

# Optimized Neural Network for Classification of Multispectral Images

Rajesh K. Agrawal<sup>a</sup>, Dr. Narendra G. Bawane<sup>b</sup>

<sup>1</sup>G. H. Raisoni College of Engineering, Department of Electronics, Nagpur, India

Email: rkdhule@yahoo.co.in

<sup>2</sup>RTM University, Nagpur, India

**Abstract**— The proposed work involves the multiobjective PSO based optimization of artificial neural network structure for the classification of multispectral satellite images. The neural network is used to classify each image pixel in various land cover types like vegetations, waterways, man-made structures and road network. It is per pixel supervised classification using spectral bands (original feature space). Use of neural network for classification requires selection of most discriminative spectral bands and determination of optimal number of nodes in hidden layer. We propose new methodology based on multiobjective particle swarm optimization (MOPSO) to determine discriminative spectral bands and the number of hidden layer node simultaneously. The result obtained using such optimized neural network is compared with that of traditional classifiers like MLC and Euclidean classifier. The performance of all classifiers is evaluated quantitatively using Xie-Beni and  $\hat{\alpha}$  indexes. The result shows the superiority of the proposed method.

**Index Terms**— Land cover classification, Multiobjective optimization (MOO), Neural network, Particle swarm optimization, Remote sensing imagery.

## I. INTRODUCTION

Multispectral images of the Earth's surface are important source of spatial data for derivation of land cover maps. We need to identify land cover class like vegetations, waterways, man-made structures and road network from satellite images. The aim of classification is to classify all pixels into one of the land cover classes. This approach is called 'per pixel' classification based on spectral data [1]. Traditional parametric statistical approaches to supervised classification are Euclidean, maximum likelihood (MLC) and Mahalanobis distance classifiers. They depend on the assumption of a multivariate Gaussian distribution for the data to be classified. But the data in feature space may not follow the assumed model. Another problem area of statistical pattern recognition in remote sensing is the "Hughes phenomenon" [2].

In recent years the ANN has been applied to general pattern recognition problems. A fundamental difference between statistical & neural approaches to classification is that statistical approach depends on an assumed model whereas neural approach depends on data [3]. In remote sensing literature, different neural network architectures are employed in supervised and unsupervised manner and for variety of purposes [4 - 8]. The neural networks have reported to yield comparable or superior accuracy compared to statistical classifiers [9]. The neural networks are particularly

suitable for remote sensing problems as they are more suitable with less reliable training samples and are less subject to "Hughes phenomenon" with properly chosen network architecture [10]. In general for supervised classification of multispectral satellite imagery, feed-forward neural network with single hidden layer is found suitable. Also the pixel grey scale value in available spectral bands is used as input feature for classification.

In our literature survey, it is found that determination of number of hidden layer neurons is critical issue and most of researchers have obtained the number of hidden layer neurons either experimentally or by same heuristics. Atkinson et al. [11] proposed that the number of hidden neuron is equal to  $[2N+1]$  where  $N$  is number of features. N. G. Kasapoglu and O. K. Ersoy [12] have empirically chosen one hidden layer with 15 neurons. A.C.Bernard, G.G. Wilkinson and I. Kanellopoulos [13] have averaged the result over tests on different neural networks with between 8 and 21 nodes in the intermediate hidden layer and found that in this range overall performance did not vary widely. S. K. Meher, B. Uma Shankar, and A. Ghosh [14] have used feed-forward MLP network fed by wavelet coefficients for IRS image classification and the nodes in the hidden layer was equal to the square root of the product of the number of input- and output-layer nodes.

In another approach, Javier Plaza et al. [15] used the MLP neural network for spectral mixture analysis. They empirically set the number of hidden layer neurons to the square root of the product of the number of input features and output classes. A. Haldera et al. [16] have used two hidden layers network for supervised classification and experimentally determined the number of hidden neurons

in each layer to get the optimum result for comparison. A Gaussian synapse artificial neural network is used by Crespo & Duro [17] to identify different crops and ground elements from remote sensing data sets. The networks are structurally adapted to the problem complexity as superfluous synapses and/or nodes are implicitly eliminated by the training procedure, thus pruning the network to the required size straight from the training set. In [18], for multispectral images, the evolved network has two hidden layer with six neurons in each layer. For networks consisting of more than one hidden layer, have not shown significant increases in accuracy compared with those containing just one.

Thus we have not found any general criteria for defining suitable network architecture. Bigger networks tend to have

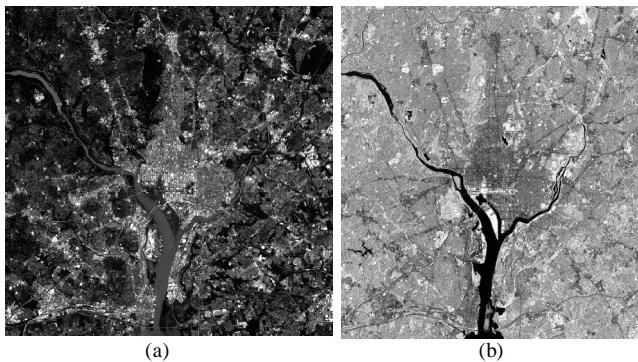


Figure 1. Multispectral images (a) band 3: visible red (b) band4: near infrared

poor generalization capability than small networks. We believe that the number of hidden layer neuron depends on the classification problem in hand and must be determined methodologically.

From our experimental analysis, it is observed that both the input feature and the number of hidden layer nodes together affect the classification accuracy and therefore must be considered simultaneously. We found that no one has work on these two issues simultaneously.

Recently, multiobjective optimization (MOO) and swarm intelligence techniques have attracted the attention of researchers in the field of satellite image processing. Y. Bazi and F. Melgani [19] proposed multiobjective PSO based method for model selection of SVMs used for satellite images. A multiobjective optimization algorithm to simultaneously optimize a number of fuzzy cluster validity indexes for classification of remotely sensed images is proposed by S. Bandyopadhyay, U. Maulik and A. Mukhopadhyay [20]. In [21], a multiobjective particle swarm optimization (MOPSO) framework is applied to estimation of the class statistical parameters and to detect discriminative bands, for clustering the hyperspectral images.

In this paper, we present MOPSO based integrated approach to find most discriminative spectral band and optimal number of hidden neuron. We also present the large number of experiments conducted to study the behavior of neural network for classification of remotely sensed imagery.

The rest of paper is organized as follows. Based on the finding during experiments, we formulate the problems associated with use of network and propose a solution on it, in section II. We briefly discuss the concept of particle swarm organization and multiobjective optimization techniques in section sections III & IV respectively. The proposed MOPSO based approach is explained in section V. Finally, result is discussed in section VI and conclusion is presented in section VII.

## II. PROBLEM FORMULATION AND SOLUTION

In this section, we describe the experiments carried out to formulate the problem associated with use of neural network for satellite image classification.

### A. Neural Network and its Topology

In our experiments we have employed single hidden layer neural network trained by back propagation algorithm [22]. The number of input nodes is determined by the number of spectral bands i.e. by dimension of the input pattern. The input pattern consists of normalize grey scale value of a pixel in selected spectral bands. Also the number of output nodes is equal to the number of classes in the image. The number of hidden layer nodes is varied from 1 to 8 for experimental analysis. After learning, the network is used as a classifier to classify the whole image.

### B. Multispectral Data

We have used Landsat satellite images of Washington DC city area [26]. The six images are of size 512 x 512 pixels each and corresponding to six spectral bands:  $b_1$ : visible blue (450 – 520 nm),  $b_2$ : visible green (520 – 600nm),  $b_3$ : visible red (630 – 670 nm),  $b_4$ : near infrared (760 – 900 nm),  $b_5$ : middle infrared (1550 – 1750 nm) &  $b_6$ : thermal infrared (10,400 – 12500 nm). The four major classes identified in the images are: water, urban area, vegetation & roads. Fig. 1 shows two images corresponding to band 3 and band 4.

### C. Training & Test Set

In our work, we have randomly selected the samples of each class by visual inspection of the image with the help of Matlab software. Total 50 samples of each class were selected and equally divided into 25 samples each to form training & test set. For training & testing input patterns, the desired output vector was obtained by setting the low value of 0.1 for the output node that do not corresponds to the pixels assigned class & high value of 0.9 for the node that does corresponds to the pixels assigned class. For example, the desired output vector for the input pattern of class 1 will be [0.9, 0.1, 0.1, 0.1], for class2, it is [0.1, 0.9, 0.1, 0.1] and so on.

### D. Experimental Framework and observations

We have performed the experiments to study the behavior of neural network for given classification problem. The numbers of spectral bands used are increased from 2 to 6. We started with spectral band combination of visible blue and red i.e. band  $b_1$  &  $b_2$ . Then we added remaining bands one by one. The number of hidden layer nodes is changed from two to eight for each of the above input feature combination. We have trained each network with training data set sizes of 5, 10 and 25 pixels. Also for each of the network, ten different initial weights were selected for training. Thus we have conducted total 1050 experiments with all combinations of above variables. Based on the result of above experiments, we made following observations.

The dependency upon initial weights can be reduced to great extent with proper input features, sufficient number of hidden nodes and adequate sample size. For proper classification, minimum number of hidden node is must. Beyond that, increase in hidden nodes does not improve the accuracy. On the contrarily, network may lose its capacity to generalize increases and the training time. The classification

accuracy is not function of the number of input features but depends upon the ‘information’ provided by the features. Therefore input features should be selected so that they contain distinct information for each output class. So we must have some method to select the useful features.

#### E. Problem statement and solution

Therefore to improve the classification accuracy and reduce computations or to increase the speed of classification, we require most discriminative spectral features and optimal number of hidden layer nodes. Thus objective is to detect most discriminative spectral band and to design an optimal ANN classifier to efficiently classify satellite images into various land cover classes.

In this work, we proposed to solve this complex problem within multiobjective particle swarm optimization framework to simultaneously estimate the most discriminative spectral band and to determine the number of nodes in hidden layer. Due to conflicting nature of both task use of multiobjective optimization is justified. The PSO based approach is employed due to its high speed of convergence.

### III. PARTICLE SWAM OPTIMIZATION

PSO is population (called as swarm) based search methodology invented by Kennedy and Eberhart [23]. It is stochastic optimization technique inspired by the social behavior of animals. Each candidate solution (particles) of a given population can benefit from its own past experiences and of all other individuals in the given population. During the iterative search process, every particle will adjust its velocity and position according to its own experience as well as those of the other particles in the swarm. Consider a swarm of size  $S$  i.e.  $P_i (i = 1, 2, \dots, S)$  and  $P_t (t)$  be the current position,  $V_i (t)$  be its velocity at iteration  $t$  and  $P_{bi} (t)$  the best position identified for  $i^{th}$  particle. Let  $P_g$  be the best global position found by the particles of the swarm. During the search process, the particles move according to the following rule [12].

$$V_i(t+1) = wV_i(t) + c_1r_1(P_{bi}(t) - P_i(t)) + \dots \\ \dots c_2r_2(P_g(t) - P_i(t)) \quad (1)$$

$$P_i(t+1) = P_i(t) + V_i(t+1) \quad (2)$$

Here  $r_1$  and  $r_2$  are random variables drawn from a uniform distribution in the range  $[0, 1]$ ,  $c_1$  and  $c_2$  are two acceleration constants with respect to the best global and local positions respectively. These parameters determine the relative weight of the self experience and the experience of group members. The inertia weight  $w$  decides tradeoff between the group and self experience capabilities of the swarm. Equation (1) allows the computation of the velocity at iteration  $(t+1)$  for each particle in the swarm and the particle position is updated with (2). These equations are iterated until maximum number of iterations is completed or the best value of the adopted fitness function is reached. Since in this application

particle have discrete binary values of 1’s and 0’s, velocity value will indicate the probability of bit taken the value 1 or 0. Therefore update formula changes to binary PSO formula as follows.

$$P_i(t+1) = 1 \text{ if } \text{sig}(V_i(t+1)) \geq 0.7 \quad (3)$$

### IV. MULTIOBJECTIVE OPTIMIZATION

In multiobjective optimization (MOO), search is performed over a number of conflicting objective functions [24]. It yields number of nondominated Pareto-optimal solutions. The aim

of search is to find the optimal vector  $\bar{x}^* = [\bar{x}_1^*, \bar{x}_2^*, \dots, \bar{x}_v^*]^T$  of  $v$  decision variables which optimizes the objective function

$\bar{f}(\bar{x}) = [f_1(\bar{x}), f_2(\bar{x}), \dots, f_k(\bar{x})]^T$  vector of  $k$  objective functions. All admissible solutions lie in the feasible region defined by the number of equality and non equality constraints. A decision vector  $\bar{x}^*$  is called Pareto optimal if and only if there is no  $\bar{x}$  that dominates  $\bar{x}^*$ . Thus  $\bar{x}^*$  is Pareto optimal if there exists no feasible vector  $\bar{x}$  which cause a reduction on some criterion without a simultaneous increase in at least another. Among the available MOO techniques, we have used, the methodology proposed by C. A. Coello Coello and M. S. Lechuga [25].

### V. PROPOSED MOPSO-ANN BASED METHOD

We shall now describe the proposed MOPSO based scheme to get subset of spectral feature and optimal number of hidden layer nodes of neural network classifier for the per pixel classification of satellite image.

#### A. Particle structure

In binary PSO each particle in the swarm is a vector that encodes the variables to be optimized in terms of binary value i.e. 1’s and 0’s.

##### 1) Input spectral bands

A part of particle encodes the candidate subset of input features among the available  $B$  spectral bands as follows.

$$\begin{aligned} f(i) &= 1 \text{ if } i^{\text{th}} \text{ spectral band is selected} \\ &= 0 \text{ if } i^{\text{th}} \text{ spectral band is not selected} \end{aligned} \quad (4)$$

where  $i = 1 \text{ to } B$

##### 2) Hidden nodes

The second part encodes the number of nodes in hidden layer as follows.

$$\begin{aligned} h(i) &= 1 & 1 \leq i \leq H \\ &= 0 & H < i \leq H_{\max} \end{aligned} \quad (5)$$

where  $H$  is selected number of nodes in the hidden layer and  $H_{\max}$  is maximum allowable nodes in the hidden layer. The structure of each particle is as shown in Fig. 2.

#### B. Fitness function

During optimization process, fitness of the particle is

evaluated by function called as fitness or objective function. The lower value indicates better fitness of the particle. In present context we need to jointly optimize the two different criteria to estimate the spectral feature and number of hidden layer nodes. The first fitness function we have used is the mean squared error (MSE) on training data set.

$$MSE = \frac{1}{N} \sum_{i=1}^N (X_i - D_i)^2 \quad (6)$$

The low MSE means less difference in desired output ( $D$ ) and actual output ( $X$ ). Hence more will be the accuracy. The MSE must be minimized to get good classification accuracy. It aims to determine the most discriminative spectral features that improve the accuracy. Thus MSE deals with our first objective to determine most discriminative features. The second objective function has to deal with the number of nodes in hidden layer and it should be in conflicting with the first fitness function, MSE. To achieve this, we proposed to use the number of nodes ( $H$ ) in hidden layer itself as a fitness parameter and it should be minimized in the optimization process. During our experimentations, we have seen that the lower the number of hidden layer nodes, more was the MSE. Thus the use of this parameter as a fitness function is justified.

### C. Algorithm description

The steps involved in the proposed algorithm for multiobjective PSO based adaption of neural network topology for pixel classification is as follows.

#### 1) Initialization

Randomly choose population size  $s$  over which search is to be performed and set the initial position of each particle  $P_i$  as follows.

- a) The number of spectral bands to be used and which band is to be selected are set randomly. Set the coordinates of selected features to 1. Keep all other coordinates to 0, as explained in (4).
- b) Randomly select the number of hidden nodes to be used i.e.  $H$  and set that number of coordinates to 1 while remaining to zero, as explained in (5).
- c) Initial velocity  $V_i(t)$  associated with the  $S$  particles is set to zero. The best position of each particle is set to its initial position, i.e.  $P_{bi} = P_g$ .

d) For each candidate particle  $P_i$ , train an ANN classifier with the encoded feature set and the number of hidden nodes. Also compute the corresponding fitness functions: MSE and  $H$ .

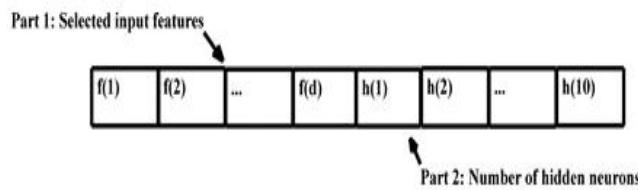


Figure 2. Structure of a PSO particle

- e) Identify the nondominated solutions by applying the algorithm described by C. A. Coello Coello and M. S. Lechuga [25] and store them in external repository.

#### 2) Search process

- a) Find the best global position from the repository and update the speed of each particle using (1).
- b) Update the position of each particle using discrete binary PSO formula (3).
- c) For each candidate particle  $P_i$  train an ANN classifier and compute the corresponding fitness functions.
- d) Identify the nondominated solutions and update the contents of repository. Also update the best position of each particle if its current position  $P_{bi}$  has a smaller fitness functions.

#### 3) Convergence:

If the termination criterion is not yet reached, return to search process.

#### 4) Classification

- a) Since best global particle represent the candidate solution having minimal cost, the spectral bands and hidden node number encoded in its structure represents the most discriminative features and optimal number of hidden layer nodes. So decode that detected spectral bands and number of hidden nodes from the structure of best global particle. Thus we get optimal neural network topology.
- b) Train such optimal network using training data set. Then use trained network in feed forward direction to classify each pixel in image.

## VI. MOPSO BASED EXPERIMENTAL RESULT AND DISCUSSION

We have implemented the proposed MOPSO algorithm on the multispectral images data set used in our experimental study described in section II. Initial parameter settings are as follows:

- Population Size= 10, 20 & 50;
- Maximum number of iteration =10, 20;
- Since we have total 6 spectral bands as input feature,  $B=6$ ;
- The maximum number of hidden layer nodes  $H_{\max} = 10$ .

We run the algorithm for number of times with different values of parameters. Each run of algorithm gives a set of

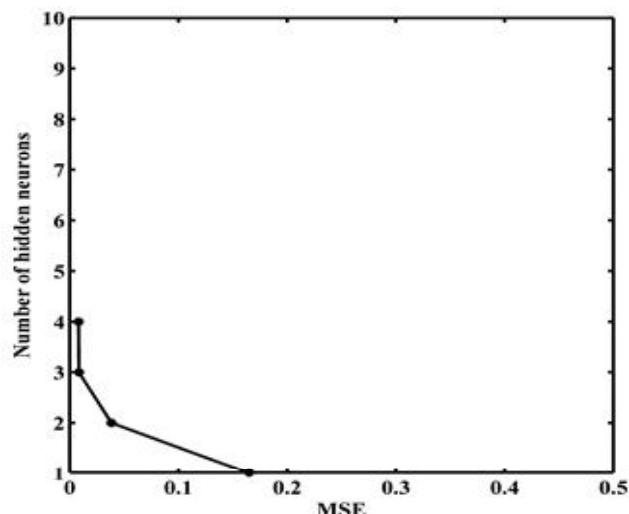


Figure 3. Pareto optimal front

nondominated solution. Fig. 3 shows such Pareto optimal front. The result of different runs of algorithm is listed in Table I. We have selected the solution having lowest MSE and minimum number of hidden neurons.

It shows that the most discriminative feature set obtained is b2, b4, b5 & b6 or b3, b4, b5 & b6. Also the optimal number of nodes in hidden layer should be three. This validates our finding in section II that for this image set the number of hidden neurons must be at least three. Thus for given classification problem, the optimal neural network structure consists of four input neurons, three hidden and four output neurons.

This neural network is trained with selected input feature pattern i.e. b3, b4, b5 & b6 and then used as classifier. The result of classification is shown in Fig. 4. All four classes are well classified and fine structures like road, bridges are also detected. Fig. 5 shows grayscale classified image. The overall test sample classification accuracy obtained was 94%.

For comparative study, the classification was also done by traditional supervised classifiers: the Euclidean classifier and maximum likelihood classifier (MLC). As shown in Table I, accuracy obtained by MLC is comparable to that obtained by our algorithm, but qualitatively classification provided by our algorithm is much better than that of MLC. MLC fails to classify finer details in the image and its accuracy varies over different runs of algorithm, due to lower number of training sample. On contrarily, even with lower sample the performance of neural classifier remains robust compared to both traditional classifiers. The XB index are 3.5, 9 and 0.75 for Euclidean, MLC, MOPSO-ANN classifier respectively. The values of  $\beta$  index are 2.2, 2 and 2.3 respectively as shown in Table II. Thus quantitatively as well as qualitatively our algorithm provides significant improvement in classification compared to both traditional classifiers.

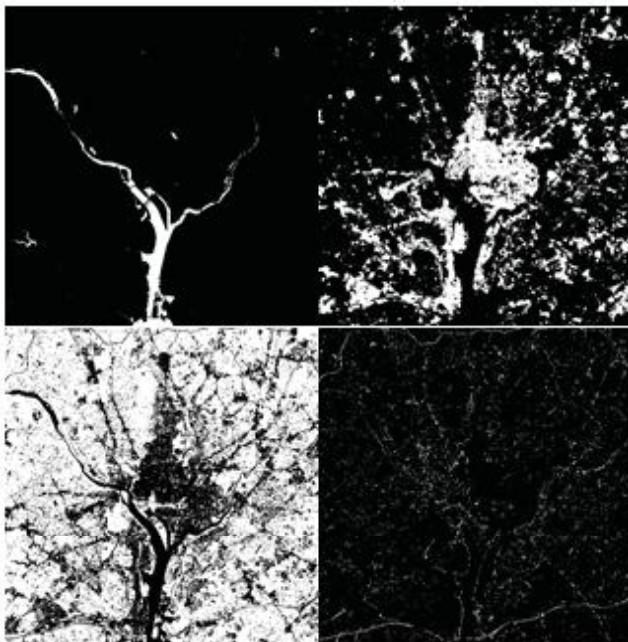


Figure 4. Classified binary images (a) river (b) urban area (c) vegetation (d) road network

TABLE I. NONDOMINATED SOLUTION AT DIFFERENT RUNS OF ALGORITHM

Iteration	Population	Detected Features	Hidden Nodes	MSE
10	10	b2b4b5b6	3	0.01
10	10	b2b4b5b6	3	0.01
20	50	b3b4b5b6	3	0.009

## VII. CONCLUSION

In this paper through our experimental study, we established that selection of most discriminative spectral bands and determination of the number of hidden layer neurons are the two most critical issues for the use of ANN in classifying the satellite images. So we presented the new methodology for efficient supervised classification of satellite image using neural network. It simultaneously estimates the most discriminative spectral features and the optimal number of nodes in hidden layer. This MOPSO based algorithm not only helps to improve the classification accuracy but also reduces the computation during classification phase of neural classifier. Our classifier is suitable for smaller number of training samples. Thus proposed work provides effective solution to the issues surfaced during our experimental study.

## ACKNOWLEDGMENT

The authors would like to thanks Rafael .C. Gonzalez, Richard .E. Woods & Steven L. Eddins for permitting us to use multispectral image data set.

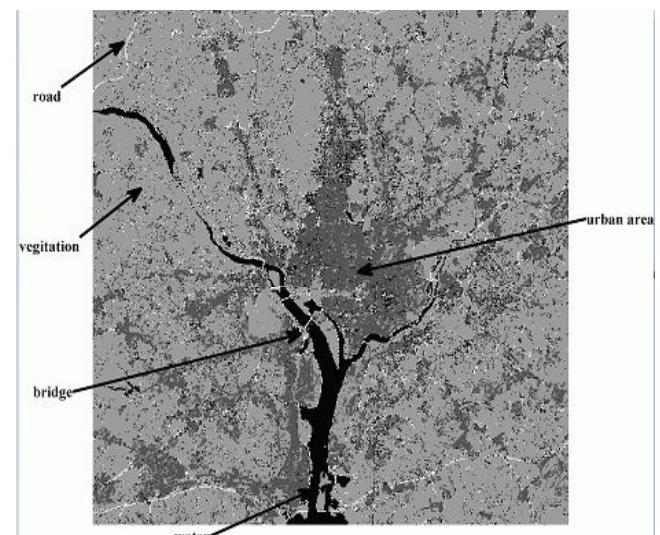


Figure 5. Grayscale classified image

TABLE II. COMPARISON WITH OTHER CLASSIFIERS

Classifier	Accuracy (Sample size:5)	Accuracy (Sample size:25)	XB index	$\beta$ index
Euclidean classifier	40%	90%	3.5	2.2
Maximum likelihood classifier (MLC)	75%	94%	9	2
MOPSO based ANN Classifier	90%	95%	0.75	2.3

## REFERENCES

- [1] P. M. Mather, Computer Processing of Remotely Sensed Images, John Wiley, UK, 1999.
- [2] D.A. Landgrebe, Signal Theory Methods in Multispectral Remote Sensing, Wiley, New York, 2003.
- [3] P. M. Atkinson, Tatnall, "Neural network in remote sensing," *Int. J. Remote Sensing*, vol. 18 (4), pp. 699-709, 1997.
- [4] J. A. Benediktsson, Swain P. H., Erosy O. K., "Neural network approaches versus statistical methods in classification of multisource remote sensing data," *IEEE Transactions on Geosciences and Remote Sensing*, vol. 28, pp. 540-552, 1990.
- [5] G. M. Foody, M. K. Arora, "An evaluation of some factors affecting the accuracy of classification by an artificial neural network," *Int. J. Remote Sensing*, vol. 18 (4), pp. 799–810, 1997.
- [6] H. M. Chee, O. K. Ersoy, "A statistical self-organizing learning system for remote sensing classification," *IEEE Transactions on Geosciences and Remote Sensing*, vol. 432 (8) (2005) pp. 1890–1900.
- [7] M. Acharyya, R. K. De, M. K. Kundu, "Segmentation of remotely sensed images using wavelet features and their evaluation in soft computing framework," *IEEE Transactions on Geosciences and Remote Sensing*, vol. 41 (12), pp. 2900-2905, 2003.
- [8] B. Mannan Joy, A. K. Ray, "Fuzzy ARTMAP supervised classification of multispectral remotely sensed images," *Int. J. Remote Sensing*, vol. 19 (4), pp. 767-774, 1998.
- [9] D. Stathakis, A. Vasilakos, "Comparison of computational intelligence based classification techniques for remotely sensed optical image classification," *IEEE Transactions on Geosciences and Remote Sensing*, vol. 44 (8), pp. 2305-2318, 2006.
- [10] C. H. Chen, P. G. Peter Ho, "Statistical pattern recognition in remote sensing," *Pattern Recognition*, vol. 41 (9), pp. 2731–2741, 2008.
- [11] P. M. Atkinson, A. R. L. Tatnall, "Neural networks in remote sensing," *Int. J. Remote Sensing*, vol. 18 (4), pp. 699-709, 1997.
- [12] N. G. Kasapoglu, O. K. Ersoy, "Border vector detection and adaptation for classification of multispectral and hyperspectral remote sensing images," *IEEE Transactions on Geosciences and Remote Sensing*, vol. 45 (12), pp. 3880 – 3893, 2007.
- [13] A. C. Bernard, G. G. Wilkinson, I. Kanellopoulos, "Training strategies for neural network soft classification of remotely sensed imagery," *Int. J. Remote Sensing*, vol. 18 (8), pp. 1851-1856, 1997.
- [14] S. K. Meher, B. Uma Shankar, A. Ghosh, "Wavelet-feature-based classifiers for multispectral remote-sensing images," *IEEE Transactions on Geosciences and Remote Sensing*, vol. 45 (6), pp. 1881–1886, 2007.
- [15] J. Plaza, A. Plaza, R. Perez, P. Martinez, "On the use of small training sets for neural network-based characterization of mixed pixels in remotely sensed hyperspectral images," *Pattern Recognition*, vol. 42, pp. 3032 – 3045, 2009.
- [16] A. Halder, A. Ghosh, S. Ghosh, "Supervised and unsupervised landuse map generation from remotely sensed images using ant based systems," *Applied Soft Computing*, vol. 11, pp. 5770–5781, 2011.
- [17] J. L. Crespo, R. J. Duro, F. Lopez, "Gaussian synapse ANNs in multi- and hyperspectral image data analysis," *IEEE Trans. Instrumentation and Measurement*, vol. 52 (3), pp. 724-732, 2003.
- [18] H. Mills, M.E.J. Cutler, D. Fairbairn, "Artificial neural networks for mapping regional-scale upland vegetation from high spatial resolution imagery," *Int. J. Remote Sensing*, vol. 27 (11), pp. 2177-2195, 2006.
- [19] Y. Bazi, F. Melgani, "Semisupervised PSO-SVM regression for biophysical parameter estimation," *IEEE Transactions on Geosciences and Remote Sensing*, vol. 45 (6), pp. 1887–1895, 2007.
- [20] S. Bandyopadhyay, U. Maulik, A. Mukhopadhyay, "Multiobjective genetic clustering for pixel classification in remotely sensing imagery," *IEEE Transactions on Geosciences and Remote Sensing*, vol. 45 (5), pp. 1506–1511, 2007.
- [21] A. Paoli, F. Melgani, E. Pasolli, "Clustering of hyperspectral images based on multiobjective particle swarm optimization," *IEEE Transactions on Geosciences and Remote Sensing*, vol. 47 (12), pp. 4175–4188, 2009.
- [22] S. Haykin, Neural Networks: A Comprehensive Foundation, Prentice Hall, England, 1997.
- [23] J. Kennedy, R. C. Eberhart, Swarm Intelligence. Morgan Kaufmann, San Mateo, 2001.
- [24] C. A. Coello Coello, M. S. Lechuga, "MOPSO: A proposal for multiple objective particle swarm optimization," in Proc. Congr. Evolutionary Computation (CEC'2002), vol. 1, pp. 1051–1056, 2002.
- [25] C. A. Coello, G. T. Pulido, M. S. Lechuga, "Handling multiple objectives with particle swarm optimization," *IEEE Trans. Evol. Comput.*, Vol. 8 (3), pp. 256–279, 2004.
- [26] R.C. Gonzalez, R. E. Woods, S. L. Eddins, Digital Image Processing using MATLAB, Pearson, Singapore, 2002.